DOT/FAA/AM-99/9

Office of Aviation Medicine
Washington, D.C. 20591

A Usability Survey of GPS Avionics Equipment: Some Preliminary Findings

Kurt M. Joseph
Civil Aeromedical Institute
Federal Aviation Administration
Oklahoma City, Oklahoma 73125

Dieter W. Jahns
SynerTech Associates
Bellingham, Washington 98227

Michael D. Nendick
The University of Newcastle
Callaghan, NSW 2308, Australia

Ross St. George
New Zealand Civil Aviation Authority
Lower Hutt, New Zealand

March 1999

Final Report

U.S. Department of Transportation
Federal Aviation Administration

NOTICE

Technical Report Documentation Page

1. Report No. DOT/FAA/AM-99/9	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle A Usability Survey of GPS Avionics Equipment: Some Preliminary Findings		5. Report Date April 1999
		6. Performing Organization Code
7. Author(s) Joseph, K.M.[1], Jahns, D.W.[2], Nendick, M.D.[3], and St. George, R.[4]		8. Performing Organization Report No.
9. Performing Organization Name and Address [1]FAA Civil Aeromedical Institute [2]SynerTech Associates	[3]The University of Newcastle [4]New Zealand Civil Aviation Authority	10. Work Unit No. (TRAIS)
		11. Contract or Grant No.
12. Sponsoring Agency name and Address Office of Aviation Medicine Federal Aviation Administration 800 Independence Ave., S.W. Washington, DC 20591		13. Type of Report and Period Covered
		14. Sponsoring Agency Code
15. Supplemental Notes		

16. Abstract

The rapid introduction of Global Positioning System (GPS) receivers for airborne navigation has outpaced the capacity of international aviation authorities to resolve human factors issues that concern safe and efficient use of such devices. Current certification technical standards appear to have had little influence on standardizing receiver architectures, interfaces, and operating manuals—despite evidence from research simulation and flight tests that lack of standardization may undermine safety. The present research used factor-analytic techniques to reduce 308 pilots' ratings from a 163-item survey to 24 factors. These factors are suitable for identifying human factors issues related to GPS receiver displays and controls, operating procedures, navigation performance, training, and other topical areas. Multivariate analysis of variance revealed that GPS experience and receiver type influenced pilot ratings for several of these factors. The results of this limited survey are consistent with previous research, and their application to certification procedures and standards is discussed.

17. Key Words GPS, Certification, Human Factors, TSO C129 A1, Pilot Experience, General Aviation, Aviation Safety	18. Distribution Statement Document is available to the public through the National Technical Information Service Springfield, Virginia 22161

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	10	

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

ACKNOWLEDGMENTS

The authors gratefully acknowledge the voluntary participation of the 308 pilots who completed survey forms. Thanks also to Ms. Tara Bergsten, Ms. Janine King, and Ms. Debbie Perry, of OMNI Corporation, for their support during survey distribution, and Dr. David Hunter, FAA Office of Aviation Medicine, for his administrative support of the survey. Finally, the authors appreciate the assistance of Mr. John Hallinan, FAA Alaskan Region Flight 2000 Program Manager, who organized survey-related activities in Alaska.

A Usability Survey of GPS Avionics Equipment: Some Preliminary Findings

INTRODUCTION

The accelerated development and introduction of Global Positioning System (GPS) receivers for use in airborne navigation has outpaced the capacity of international aviation authorities to fully implement regulations and guidance for the safe and efficient use of such devices (Nendick & St. George, 1996a). Technical Standard Order (TSO) C129 A1 is the document currently used to certify standalone, Instrument Flight Rules (IFR) GPS receivers for installation in the United States and Canada, and is accepted as the certification standard elsewhere, including Australia and New Zealand. However, it appears to have had little influence on standardizing receiver architectures, interfaces, and operating manuals (Heron, Krolak, & Coyle, 1997). Recent flight simulation research performed at the Federal Aviation Administration Civil Aeromedical Institute (FAA/CAMI) illustrates that some features of GPS-receiver interfaces can compromise their effective use and perhaps undermine safety (Williams, 1998; Wreggit & Marsh, 1998). Furthermore, FAA flight tests using brand-name receivers have revealed several human factors issues related to pilot interaction with receiver interfaces (Williams, in press; Winter & Jackson, 1996). These reports underscore the need for GPS-receiver standards that exceed the criteria specified in TSO C129 A1. Human factors issues associated with receiver architectures, interfaces, and operating manuals need to be specifically addressed.

At least two human factors references directly support TSO C129 A1. McAnulty's (1994) review of human factors principles and guidelines for the design of controls and displays for standalone GPS and Long Range Navigation (LORAN) receivers is germane to regulatory requirements. In addition, the FAA has issued an aircraft certification human factors and operations checklist for standalone GPS receivers (Huntley, Turner, Donovan, & Madigan, 1995). It includes both a bench and a flight test, which were designed to assist certification personnel and manufacturers in evaluating receiver interface characteristics of GPS receivers to be certified in accordance with TSO C129 A1. The bench test assists in evaluations of GPS receivers that do not require aircraft installation and a test flight. The flight test helps in evaluating GPS-receiver design characteristics and functions during actual flight conditions. Both checklist tests focus on GPS-receiver controls, displays, and operating characteristics. Together with the guidelines proposed by McAnulty (1994), the checklist tests represent important progress in resolving significant human factors issues associated with GPS receivers and their certification. However, evidence that considerable work remains on these issues is manifest in requests from GPS manufacturer and certification personnel for guidance on numerous human factors issues for which they have no baseline data.

The research reported here is intended to complement existing human factors data by providing baseline measures of pilot perceptions of, and experiences with, GPS receivers. This research is based on the 163-item GPS User Survey 97 (Nendick, 1997), which was administered to 1,880 pilots in the United States (US). Nendick (1994) administered an earlier version of this survey to 227 New Zealand pilots and is currently administering a further-modified, 1998 version to 4,000 Australian and New Zealand pilots. In the US, the survey represents the first systematic attempt to quantify a large sample of pilot perceptions of GPS receivers. Together with the survey data from New Zealand and that forthcoming from Australia, the US data provide a unique source of information that can be aggregated with existing human factors data (e.g., usability tests) to properly support the tasks of pilots, certification personnel, and GPS-receiver manufacturers. One of the many challenges for the FAA as it moves toward greater dependence on GPS-based navigation is to supplement TSO C129 A1 and future technical standard orders with human factors specifications. These should ensure usability through some basic standardization guidelines, without prohibiting GPS-receiver manufacturers from adding new features to their GPS devices.

METHOD

Administration of the GPS User Survey was part of a larger effort to explore free-flight enabling technologies, like GPS, and their impact on pilot performance of tasks related to communication, navigation, and surveillance functions. During the course of its modification for use in the US, the original 125-item survey developed by Nendick (1994) was expanded to include 163 questions. These questions were categorized into the following topic areas: Receiver Controls and Displays, Operating Logic, Receiver Functions, Receiver Operations, Operating Procedures, Navigation Performance, Pilot Attitudes, and Training. Seven-point, Likert-type rating scales were used to measure pilots' responses to survey questions. Several open-ended questions were included within the survey to gather qualitative data on particular issues that could not be addressed using the question and rating format. Finally, several demographic questions were inserted to determine the extent to which the sample of pilots used in this research represented the General Aviation (GA) pilot population.

The 35-minute survey was administered to pilots in two different FAA Regions. Pilots had to have a current medical certificate, be at least a private pilot, and have a minimum of five hours of GPS flight time to be eligible for participation in the survey. Eighty-eight pilots from the FAA Southwest Region (ASW) completed the survey at FAA/CAMI in December 1997. These pilots were part of a group of 150 pilots who participated in several research studies during a one-day session at FAA/CAMI. Thus, nearly 59% of the pilots in the ASW Region sample completed the survey. In April 1998, the survey was mailed to 1,800 FAA Alaskan Region (AAL) pilots, of which approximately 12% (N=220) completed it. This low return rate is consistent with that obtained for other surveys of AAL Region pilots (e.g., see Driskill, Weissmuller, Quebe, Hand, & Hunter, 1997).

The average age of pilots in the AAL and ASW Regions was 49 (SD=10) and 41 (SD=14), respectively. Table 1 shows the average, median, and standard deviation for total hours and GPS experience (in hours) for each pilot sample. Pilots from both regions were comparable in total hours. A comparison of GPS experience revealed that AAL Region pilots had nearly twice as many GPS hours than pilots in the ASW Region. The 50th percentile for GPS experience was 200 hours.

Table 2 compares the actual distribution of licenses within each region (see 1996 FAA Statistical Handbook of Aviation) with the distribution obtained for AAL and ASW Region samples. Observed percentages for each region sum to less than 100% because a small percentage of pilots did not indicate which license they held. Although the AAL Region sample is skewed more heavily, analysis of the license distribution in both regions indicates that private pilots comprised a very large majority of all survey respondents. Given the skewed distribution and low return rate, the results reported in the next section would be most representative of GA pilots who hold a private pilot license.

RESULTS

Estimates of internal consistency (i.e., coefficient alpha) were calculated for each of the eight survey topic areas. In general, the following standardized

Table 1. Demographics for Pilot Samples

FAA Region	N	Total Hours			GPS Hours		
		Mean	Med	SD	Mean	Med	SD
AAL	220	1883	988	2919	414	240	621
ASW	88	1999	957	2887	210	40	448
Both	**308**	**1917**	**983**	**2905**	**354**	**200**	**582**

Table 2. Pilot License as a Function of Region

FAA Region	License Type (Actual and Observed in %)					
	Private		Commercial		Air Transport	
	Act.	Obs.	Act.	Obs.	Act.	Obs.
AAL	49.7	89.1	29.2	1.8	21.1	5.0
ASW	46.0	37.5	25.9	38.6	28.1	19.3

item alphas indicate that pilot ratings for most of the survey topic areas were reliable: Receiver Controls and Displays (.92), Operating Logic (.90), Receiver Functions (.89), Receiver Operations (.91), Operating Procedures (.60), Navigation Performance (.70), Pilot Attitudes (.81), and Training (.56). The relatively low reliability estimates for Operating Procedures and Training may be due to the smaller number of items in these topic areas and the large variety of GPS receivers, each of which has a different operating manual.

Underlying factor structures were extracted by entering pilot ratings from each of the survey topic areas into separate exploratory principal components factor analyses that utilized varimax rotation to obtain factor solutions. A multivariate analysis of variance (MANOVA) was then utilized to determine if differences in mean factor ratings were a function of FAA region (AAL and ASW), GPS experience (below the 50th percentile and above the 50th percentile), and GPS-receiver type (hand-held and panel-mount). Pilots with 200 hours or less ranked below the 50th percentile for GPS experience, whereas pilots with more than 200 hours ranked above the 50th percentile.

Pilot ratings were pooled across FAA region after preliminary analyses revealed differences that were attributable solely to the availability of ground-based navigational aids within each region. For example, regions located in the contiguous US have more very high frequency, omni-directional range (VOR) navigational aids than the AAL Region. Consequently, pilots in the AAL Region were less likely to use a VOR navigational aid as a backup to GPS than were pilots in the ASW Region. Pilots in the AAL Region also were less likely to use a backup means of navigation because this region has fewer navigational aids than do regions located in the contiguous US. Hence, the following results are based on the exploratory factor analyses of each survey topic area and the MANOVA, which included GPS experience and GPS-receiver type as independent variables, and mean factor ratings within each survey topic area as dependent variables.

Receiver Controls and Displays

Consistent with the findings of Nendick (1994), the following four factors emerged from the factor analysis of this topic area.

- Control Dimensions (button and/or knob operation, labeling, and layout)
- Accessories and Installation (power supply, antenna, receiver size and shape, fit into aircraft)
- Display Readability (readability of symbols and day/night illumination)
- Display Messages (warning indications)

Generally, higher ratings on these factors indicate a favorable view of the particular receiver that was rated. MANOVA tests of pilot ratings revealed that the main effect of GPS-receiver type was significant for both the Accessories and Installation and the Display Readability factors. Relative to hand-held receivers (M=5.6, SD=1.0), the accessories and installation associated with panel-mount receivers (M=6.0, SD=0.8) were assigned significantly higher ratings by pilots [$F(1,184)=6.8$, $p<.01$]. Panel-mount receivers (M=5.4, SD=1.1) also received significantly higher ratings on the Display Readability factor than did hand-held receivers [M=6.0, SD=1.0; $F(1,184)=11.2$, $p<.01$]. Despite these differences, all pilots gave fairly high ratings to the four factors that characterized receiver controls and displays.

Operating Logic

Two factors emerged from the factor analysis of this topic area.

- Programming Demands (route programming and review, waypoint entry, standardization)
- Cognitive Demands (practice required for proficiency, reliance on memory, operating complexity)

None of the interaction and main effect tests involving these factors was significant.

Receiver Functions

Two factors emerged from the factor analysis of this survey topic area.

- Navigation Feedback (desired track, distance to go, crosstrack error and CDI, and groundspeed)
- Receiver Integrity and Emergency Information (RAIM warning, satellite status, nearest airport, and present position)

None of the interaction and main effect tests involving these factors was significant.

Receiver Operations

Two factors emerged from the factor analysis of this survey topic area.

- •In-flight Operation (one-handed operation, use in turbulence, and night and day use)
- •Data Entry and Modification (creating and modifying route, pre-flight and in-flight data entry, undoing errors)

In general, higher ratings on these factors indicate a favorable impression of the particular receiver that was rated. There was a significant effect of GPS experience for the In-flight Operation factor [$F(1,184)=4.0$, $p<.05$]. Pilots with more GPS experience assigned significantly higher ratings ($M=5.7$, $SD=1.1$) for in-flight operation of receivers than pilots with less GPS experience ($M=5.3$, $SD=1.1$). Once again, the overall ratings for these factors were fairly high.

Operating Procedures

Four factors emerged from the factor analysis of this survey topic area.

- •Validation of Receiver Data (in-flight cross-checks of information using other data sources)
- •Operational Errors (incorrect data input, misreading the display, and mode errors)
- •Navigation Contingencies (navigation backups for GPS, use of other navigational methods with GPS)
- •Incidents (incidents resulting from misreading the display or incorrect input of data)

A significant main effect test for the Operational Errors factor revealed that the amount of GPS experience influenced pilot ratings [$F(1,184)=5.0$, $p<.03$]. Pilots with less GPS experience assigned significantly higher ratings ($M=2.8$, $SD=1.0$) for this factor than did pilots with more GPS experience ($M=2.5$, $SD=1.0$). This effect suggests that pilots with less experience report that they make more errors during their interaction with GPS receivers.

Navigation Performance

Four factors emerged from the factor analysis of this survey topic area.

- •Awareness Issues (changes in situation awareness, lookout, and instrument scan)
- •Workload Issues (changes in mental and physical workload, head-down time, and chart use)
- •Course Tracking (tracking direct and to waypoints, and accuracy of tracking)
- •Decision Making (navigation decisions, flying in marginal conditions, and flying near controlled airspace)

The main effect of GPS experience was significant for the Course Tracking factor [$F(1,184)=4.8$, $p<.03$]. Pilots with more GPS experience assigned significantly higher ratings ($M=5.9$, $SD=0.9$) for this factor than did pilots with less GPS experience ($M=5.5$, $SD=1.0$). This finding suggests that pilots perceive improvements in GPS course tracking as concomitant with gains in GPS experience.

Pilot Attitudes

Four factors emerged from the factor analysis of this survey topic area. The first two factors are consistent with the findings of Nendick (1994).

- •Confidence in GPS (use for IFR navigation and non-precision approaches, and accuracy and reliability of GPS device)
- •User Confidence (use of basic and complex GPS functions)
- •Dependence on GPS (complacency, reliance on GPS for VFR and IFR)
- •Use of Database (affordability and use of current data)

The main effect of GPS experience was significant for the User Confidence and Dependence on GPS factors. [$F(1,184)=5.6$, $p<.02$; $F(1,184)=4.5$, $p<.04$, respectively]. Pilots with more GPS experience assigned significantly higher ratings ($M=6.0$, $SD=0.8$) for the User Confidence factor than did pilots with less GPS experience ($M=5.6$, $SD=1.0$). The average ratings for this factor were very high. Compared with those who had less experience ($M=3.9$, $SD=1.2$), pilots with more experience also assigned significantly higher ratings ($M=4.3$, $SD=1.3$) for the Dependence on GPS factor. Overall, the ratings indicated that pilots reported increased dependence on GPS as they gain more experience with it. The main effect test of GPS-receiver type for the Dependence on GPS factor approached the nominal significance level criterion of .05 [$F(1,184)=2.9$, $p<.09$]. Pilots assigned higher ratings to panel-mount receivers ($M=4.4$, $SD=1.2$) than they did to hand-held receivers ($M=3.9$, $SD=1.2$), although this difference is not statistically significant.

Training

Two factors emerged from the factor analysis of this survey topic area.

- •Knowledge and Experience Required (technical training, and knowledge level and human factors training needed)
- •Knowledge and Experience Attained (user knowledge, training received, and user manual)

The main effect of GPS-receiver type was significant for the Knowledge and Experience Required factor [$F(1,184)=10.1$, $p<.01$]. Pilots assigned higher ratings to panel-mount receivers ($M=4.5$, $SD=1.0$) than they did to hand-held receivers ($M=4.0$, $SD=1.1$). This finding indicates pilots perceive that panel-mount receivers require more training to operate proficiently. Finally, the main effect test of GPS experience for the Knowledge and Experience Attained factor approached the nominal significance level criterion of .05 [$F(1,184)=3.0$, $p<.09$]. Pilots with more GPS experience assigned higher ratings ($M=4.4$, $SD=1.1$) than did pilots with less GPS experience ($M=4.2$, $SD=1.0$), although this difference is not statistically significant.

DISCUSSION

The present research, which is based on a limited survey, has produced several meaningful results. First, MANOVA tests indicated that GPS experience and GPS-receiver type did not interact to affect pilot factor ratings within each of the survey topic areas. Thus, the following discussion assumes that GPS experience and receiver type combine in an additive fashion to influence pilot perceptions of GPS receivers. Significant differences in pilot ratings as a function of GPS-receiver type suggested that panel-mount receivers were better than hand-held receivers on factors such as display readability, power supply, antenna, size and shape, and fit into aircraft. Pilot ratings also revealed that panel-mount receivers require more knowledge and training to operate proficiently. This finding was not statistically significant, although it is of practical importance. So, too, is the statistically non-significant finding that pilots who used panel-mount receivers rated themselves as more complacent and reliant on GPS than pilots who used hand-held receivers. In part, this finding may be because panel-mount receivers provide continuous monitoring of GPS signal integrity, whereas hand-held receivers do not.

Significant differences in pilot ratings as a result of GPS experience were distributed across several survey topic areas. Pilot ratings suggested that, with more experience, in-flight operation of GPS receivers became easier, fewer operational errors were made, course tracking improved, there were gains in knowledge of the receiver, and increased confidence in using basic and complex functions. One possible negative consequence associated with more GPS experience is that pilots perceive themselves to be more reliant on GPS and more complacent than pilots with less experience.

Consistent with ratings of the New Zealand pilots surveyed by Nendick (1994), the results reported here provide evidence of a discrepancy between pilot ratings of GPS-receiver design and pilot performance in GPS-receiver evaluations (e.g., see Wreggit & Marsh, 1998). A plausible explanation for this discrepancy is presented by Clarke (1994) and Nendick and St. George (1996b), who have warned of the pitfalls related to pilot perceptions of the operational simplicity of GPS. On one hand, pilots may be captivated by the inherent simplicity and minimal training required to execute frequently used basic receiver functions. On the other hand, as Winter and Jackson (1995) have demonstrated, pilots who have not acquired a significant amount of knowledge and training can become overwhelmed when they go beyond basic receiver functions to complex ones that are required for difficult navigational tasks. Examples of subtle human factors issues associated with complex receiver functions will likely be revealed during flight tests using GPS Wide Area Augmentation System (WAAS) terminal approaches and in future tests of other difficult navigational tasks.

Another meaningful finding resulted from the exploratory factor analysis of each survey topic area. Separate sets of factors were identified that could be used as a basis for supplementing existing certification guidelines for GPS receivers. In combination, certification personnel, manufacturers, and human factors specialists could examine the specific factors within each topic area for design issues relevant to certification. Such an effort would result in coverage of traditional human factors issues related to receiver displays and controls, as well as other topical issues, such as those related to operating procedures, navigation performance, and training.

This approach to identifying human factors issues in certification has at least two advantages over the current approach. The first advantage is that certification personnel and manufacturers would collaborate with human factors specialists in examining issues and developing materials. Just as it is for GPS receivers, end-user participation is critical for the acceptance and use of certification test materials. Second, standardization of tests, checklists, and other certification material will go a long way to ensure that GPS receivers have standardized architectures, interfaces, and operating manuals. Current practices place considerable onus on certification personnel to gather information and develop their own supplemental materials for certification tests. Consequently, subtle variations in certification materials and procedures impose greater demands on certification personnel and, perhaps, make the task of GPS-receiver standardization more difficult. Adopting a systematized approach for developing human factors standards in support of certification personnel and GPS manufacturers may ease this difficulty.

REFERENCES

Clarke, B. (1994). *Aviator's guide to GPS.* New York, NY: TAB Books.

Driskill, W.E, Weissmuller, J.J., Quebe, J.C., Hand, D.K., & Hunter, D. (1997). The use of weather information in aeronautical decision-making: II. U.S. Department of Transportation, Federal Aviation Administration, Office of Aviation Medicine, Washington, D.C. DOT/FAA/AM-97/23. NTIS # ADA340406.

FAA (No date). *FAA Statistical Handbook of Aviation*, pp. 7-4/7-9 (Online). Available: <http://api.hq.faa.gov/handbook/1996/toc96.htm> (1998, July 28).

Heron, R., Krolak, W. & Coyle, S. (1997). *A human factors approach to the use of GPS receivers.* Vancouver, BC: Heron Ergonomics, Incorporated.

Huntley, Jr., M.S., Turner, J.W., Donovan, C.S., & Madigan, E. (1995). *FAA aircraft certification human factors and operations checklist for standalone GPS receivers (TSO C129 Class A).* Technical Report DOT/FAA/AAR-95/03. Washington, DC: Federal Aviation Administration.

McAnulty, D.M. (1994). *Guidelines for the design of GPS and LORAN receiver controls and displays.* Technical Report DOT/FAA/RD-95/01. Washington, DC: Federal Aviation Administration.

Nendick, M.D. (1994). *Global Positioning System (GPS): Human factors aspects for general aviation pilots.* Unpublished masters thesis. Palmerston North, New Zealand: Massey University.

Nendick, M.D. (1997). *GPS User Survey 97.* Callaghan, NSW, Australia: The University of Newcastle.

Nendick, M.D. & St. George, R. (1996a). General aviation pilots and GPS: Some results from a New Zealand study. *Airways, 9*, 12-6.

Nendick, M.D. & St. George, R. (1996b). Human factors aspects of global positioning systems (GPS) equipment: A study with New Zealand pilots. In R. Jensen (Ed.) *Proceedings of the Eighth International Symposium on Aviation Psychology*, pp. 152-7. Columbus, OH: The Ohio State University.

Williams, K.W. (1998). *GPS design considerations: Displaying nearest airport information.* U.S. Department of Transportation, Federal Aviation Administration, Office of Aviation Medicine, Washington, D.C. DOT/FAA/AM-98/12. NTIS # ADA346043.

Williams, K.W. (in press). *GPS user-interface design problems.* U.S. Department of Transportation, Federal Aviation Administration, Office of Aviation Medicine, Washington, D.C.

Winter, S. & Jackson, S. (1996). *GPS issues.* Technical Report DOT/FAA/AFS450. Oklahoma City, OK: Federal Aviation Administration.

Wreggit, S. & Marsh III, D. (1998). *Cockpit integration of GPS: Initial assessment- menu, formats and procedures.* U.S. Department of Transportation, Federal Aviation Administration, Office of Aviation Medicine, Washington, D.C. DOT/FAA/AM-98/09. NTIS # ADA341122.